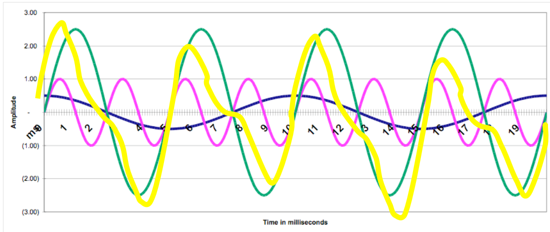
Sounds of English

Topic 8
The acoustics of speech:
Amplitude and Spectrograms

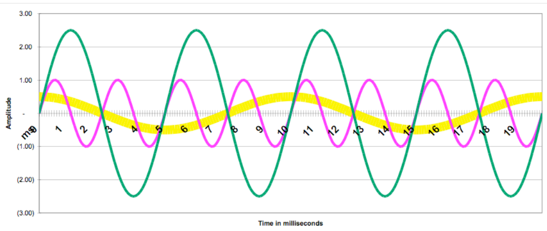
Another way to look at waves

- Last time we saw that three simple waves with different amplitudes and frequencies can combine to make a complex wave:



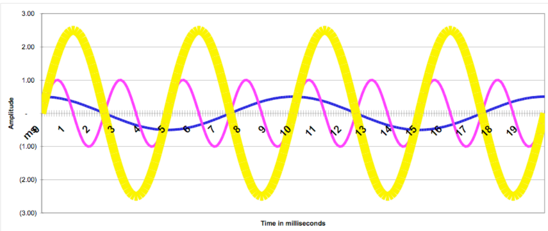
Another way to look at waves

- And we were able to measure the amplitude and frequency of each wave:
  wave 1: Amp = 0.5, f = 100



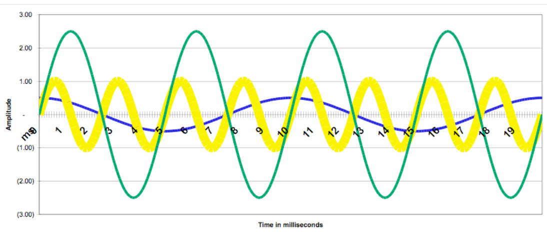
Another way to look at waves

- And we were able to measure the amplitude and frequency of each wave:
  wave 2: Amp = 2.5, f = 200



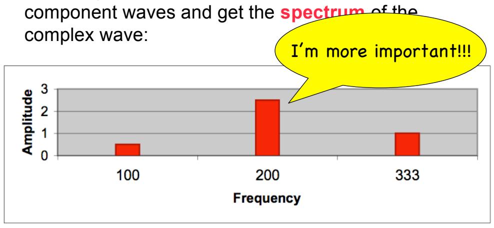
Another way to look at waves

- And we were able to measure the amplitude and frequency of each wave:
  wave 3: Amp = 1, f = 333



The spectrum

- We can graph the frequencies and amplitudes of the component waves and get the **spectrum** of the complex wave:



note that the amplitudes are different - waves at certain frequencies are "more important" than others

The source-filter method

- Waves generated by vibrating vocal folds can get even more complicated!
- And the vocal tract is a very complex tube
- Depending on the shape of our vocal tract
  - some frequencies of periodic waves will be amplified
  - other frequencies will be damped

the vocal tract acts as "filter" - it picks out certain parts of the "source" sound from our vibrating vocal folds and makes them more prominent

More about amplitude

- Measured in deciBels (dB)
- Not an absolute value, but a **ratio**
- 1 dB =
  - 10 times the log of the ratio of 2 sounds' intensities
- Example
  - if sound A is 30 dB louder than sound B,
  - 10 times the log of the ratio = 30
  - the log of the ratio = 3 (because 10 x 3 = 30)
  - so the ratio of sound A to sound B is 10^3 : 1
  - 10^3 = 1000 (a 1 followed by 3 zeros)
  - so sound A is 1000 times louder than sound B !

Another example of deciBels

- We use a log scale is because the differences between sounds can get REALLY big:
  - if sound A is 120 dB louder than sound B,
  - 10 times the log of the ratio = 120
  - the log of the ratio = 12 (10 x 12 = 120)
  - so the ratio of sound A to sound B is 10^{12} : 1
  - 10^{12} = 1,000,000,000,000 (a 1 followed by 12 zeros)
  - so sound A is a million million times louder than sound B !

dB scale of common sounds

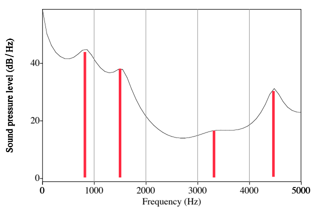
Intensity (dB)	Sound
130	4-engined jet aircraft, 120 ft away
120	threshold of pain; pneumatic hammer 3 ft away
110	boilermakers' shop, 'rock' band
100	car horn 15 ft away; symphony orchestra *fortissimo*
90	pneumatic drill, 4 ft away
80	noisy tube train; loud radio m
70	telephone bell, 10 ft away
60	conversation, 3 ft away; car, 30 ft away
50	quiet office
40	residential area, no traffic; subdued conversation
30	quiet garden; whispered conversation
20	ticking of watch; broadcast studio
10	rustle of leaves
0	threshold of audibility

maximum volume of your ipod

routine exposure to over 85 dB can produce hearing loss over time

Back to amplitude & frequency

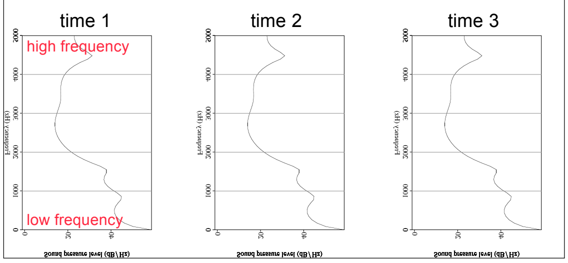
- The spectrum below shows the resonating (or amplified) frequencies of me saying a vowel at just one moment in time



which are the resonant frequencies?

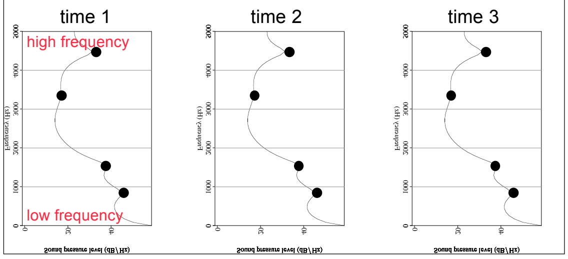
Making a spectrogram

- Instead of looking at just one moment in time, what if we rotated the spectrum on edge and looked at a whole bunch of them?



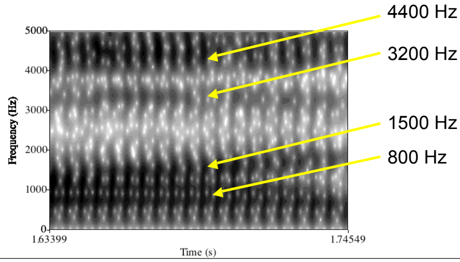
Making a spectrogram

- And now what if we made the “peaks” really dark and turned each spectrum on its edge (so that the peaks would be pointing directly at us)...



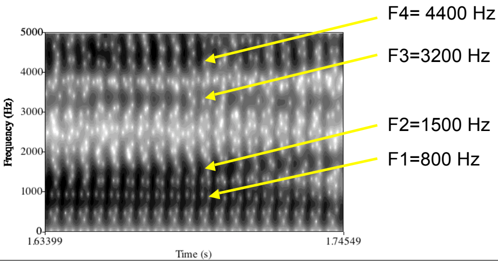
We would get...

- ...a spectrogram!
- Notice the dark lines at the same frequencies where the “peaks” were in the spectrum



Formants

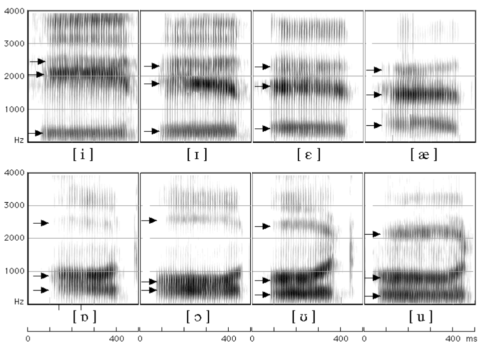
- The dark lines represent resonant frequencies
- We call them FORMANTS



Why do we care?

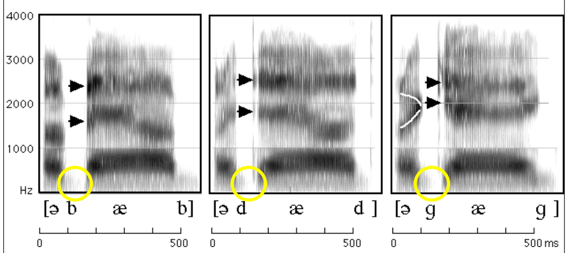
- Phoneticians use speech spectrograms to look at sounds
- Different speech sounds will have different spectrograms
  - because the shape of the vocal tract will be different for various speech sounds
  - so the frequencies that are allowed to “resonate” (or be amplified) will be different
- Some sounds will have formants, others won’t
  - vowels and sonorant consonants (i.e., periodic sounds) will have formants
  - noisy consonants like plosives and fricatives will not

Vowel spectrograms



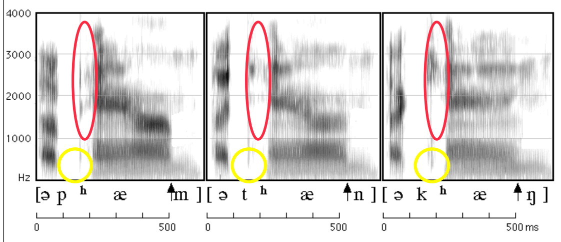
adult male speaker (British English) - Ladefoged 2001

Voiced plosive spectrograms



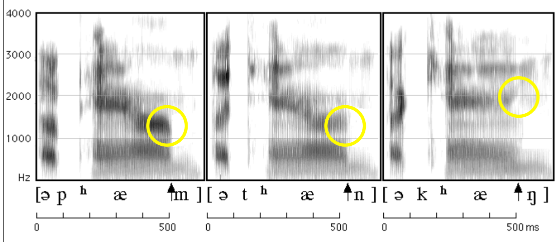
notice the voicing “bar” at the very bottom
this is the fundamental frequency
(vibrating vocal folds)

Voiceless plosive spectrograms



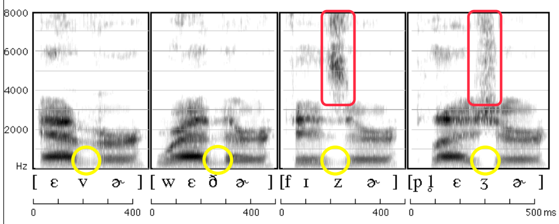
NO voicing “bar” at the bottom
but a very noisy release burst after the closure!

Nasal spectrograms



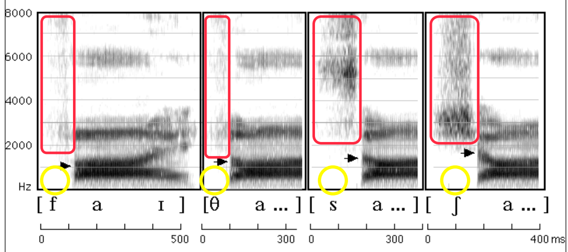
Notice how the transitions between the [æ] and
the following nasal are different depending on
place of articulation of the nasal

Voiced fricative spectrograms



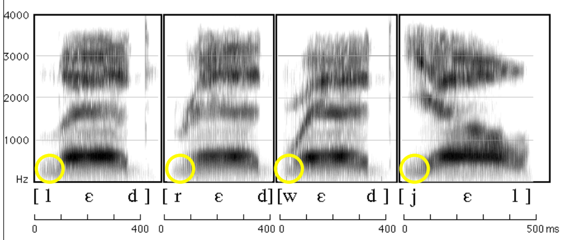
Notice:
voicing bar
formants and their transitions
high frequency noise for [z] and [ʒ]

Voiceless fricative spectrograms



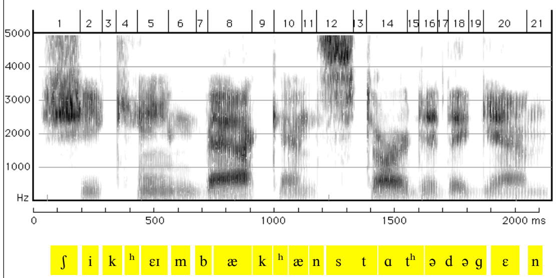
Notice:
NO voicing bar
no formants, but different transitions to vowels
lots of high frequency noise for [s] and [ʃ]

Approximant spectrograms



Notice:
voicing bar
formants and transitions to following vowel

An English sentence



“She came back and started again.”

References

Spectrograms from:

Ladefoged, Peter. 2000. *A Course in Phonetics* (5th Ed.). Thomson/Wadsworth

Also available at the UCLA Phonetics Lab Data site at:

http://www.phonetics.ucla.edu/course/chapter8/figure8.html

Practice with Praat

- Reading files
- Opening files in the editing window
- Zooming in and out
- Showing the spectrogram
- Selecting parts of the waveform/spectrogram
- Extracting and drawing parts of a sound

We're done!!

See you all next week.

Same bat time.

Same bat channel.